

Title (en)

Method of speech recognition using variational inference with switching state space models

Title (de)

Spracherkennungsverfahren das Variationsinferenz mit veränderlichen Zustandsraummodellen benutzt

Title (fr)

Méthode pour la reconnaissance de parole utilisant de l'inférence variationnelle avec des modèles d'espace d'état changeants

Publication

EP 1465154 A3 20070606 (EN)

Application

EP 04007985 A 20040401

Priority

US 40516603 A 20030401

Abstract (en)

[origin: EP1465154A2] A method is developed which includes 1) defining a switching state space model for a continuous valued hidden production-related parameter and the observed speech acoustics, and 2) approximating a posterior probability that provides the likelihood of a sequence of the hidden production-related parameters and a sequence of speech units based on a sequence of observed input values. In approximating the posterior probability, the boundaries of the speech units are not fixed but are optimally determined. Under one embodiment, a mixture of Gaussian approximation is used. In another embodiment, an HMM posterior approximation is used.

IPC 8 full level

G10L 15/14 (2006.01); **G10L 15/06** (2013.01); **G10L 15/10** (2006.01)

CPC (source: EP KR US)

B01D 21/2433 (2013.01 - KR); **B01D 21/245** (2013.01 - KR); **C02F 1/40** (2013.01 - KR); **G10L 15/14** (2013.01 - EP US)

Citation (search report)

- [PX] US 6591146 B1 20030708 - PAVLOVIC VLADIMIR [US], et al
- [PX] US 6678658 B1 20040113 - HOGDEN JOHN [US], et al
- [X] GHAHRAMANI Z ET AL: "VARIATIONAL LEARNING FOR SWITCHING STATE-SPACE MODELS", NEURAL COMPUTATION, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, US, vol. 12, no. 4, April 2000 (2000-04-01), pages 831 - 864, XP009046561, ISSN: 0899-7667
- [X] BRIDLE J S ET AL: "An Investigation of Segmental Hidden Dynamic Models of Speech Coarticulation for Automatic Speech Recognition", INTERNET CITATION, 1998, XP002318768, Retrieved from the Internet <URL:http://www.clsp.jhu.edu/ws98/projects/dynamic/presentations/final/clsp.p> [retrieved on 20050221]
- [A] MA J ET AL: "A mixture linear model with target-directed dynamics for spontaneous speech recognition", 2002 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS. (ICASSP). ORLANDO, FL, MAY 13 - 17, 2002, IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP), NEW YORK, NY : IEEE, US, vol. VOL. 4 OF 4, 13 May 2002 (2002-05-13), pages I - 961, XP010804855, ISBN: 0-7803-7402-9
- [A] DENG LI ET AL: "Spontaneous speech recognition using a statistical coarticulatory model for the vocal-tract-resonance dynamics", JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, AIP / ACOUSTICAL SOCIETY OF AMERICA, MELVILLE, NY, US, vol. 108, no. 6, December 2000 (2000-12-01), pages 3036 - 3048, XP012002060, ISSN: 0001-4966
- [PX] LEE L J ET AL: "Variational inference and learning for segmental switching state space models of hidden speech dynamics", 2003 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS. (ICASSP). HONG KONG, APRIL 6 - 10, 2003, IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP), NEW YORK, NY : IEEE, US, vol. VOL. 1 OF 6, 6 April 2003 (2003-04-06), pages I - 872, XP002331640, ISBN: 0-7803-7663-3

Designated contracting state (EPC)

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LI LU MC NL PL PT RO SE SI SK TR

Designated extension state (EPC)

AL HR LT LV MK

DOCDB simple family (publication)

EP 1465154 A2 20041006; **EP 1465154 A3 20070606**; **EP 1465154 B1 20091014**; AT E445896 T1 20091015; CN 1534597 A 20041006; DE 602004023555 D1 20091126; JP 2004310098 A 20041104; KR 20040088368 A 20041016; US 2004199386 A1 20041007; US 2005119887 A1 20050602; US 6931374 B2 20050816; US 7487087 B2 20090203

DOCDB simple family (application)

EP 04007985 A 20040401; AT 04007985 T 20040401; CN 200410032697 A 20040331; DE 602004023555 T 20040401; JP 2004109419 A 20040401; KR 20040022168 A 20040331; US 40516603 A 20030401; US 98460904 A 20041109